

LOAD DISTRIBUTOR

Publication number: JP2004094555

Publication date: 2004-03-25

Inventor: KIKUCHI SHUNSUKE; OGAWA ATSUSHI

Applicant: FUJITSU LTD

Classification:

- international: G06F13/00; G06F15/16; H04L1/16; H04L12/56; H04L29/08; G06F13/00; G06F15/16; H04L1/16; H04L12/56; H04L29/08; (IPC1-7): G06F13/00; H04L1/16; H04L12/56

- European: H04L29/08N9A

Application number: JP20020254233 20020830

Priority number(s): JP20020254233 20020830

Also published as:

US2004054796 (A1)

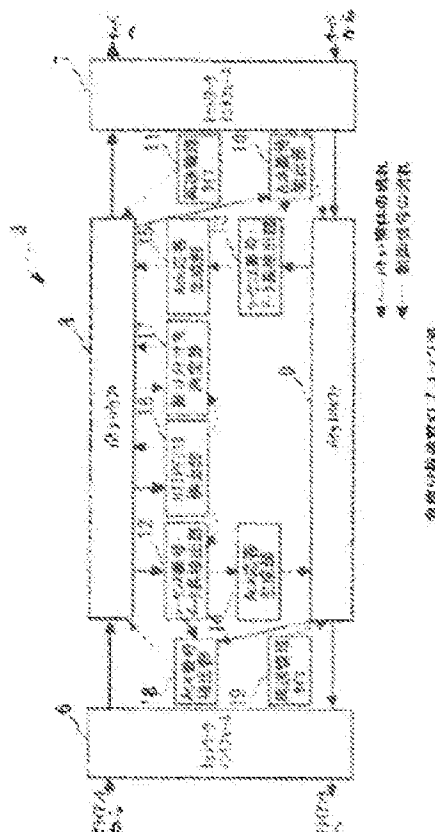
Report a data error here

Abstract of JP2004094555

PROBLEM TO BE SOLVED: To eliminate retransmission control for a request received from a client to reduce a use time of a transmission buffer.

SOLUTION: This distributor is provided with a determination means for determining the propriety of transmitting a confirmation reply for a data to the client in response to a content of the data received from the client, a confirmation responding means for transmitting a confirmation reply to the client in response to a determined result in the determination means, a transfer means for transferring a data received from the client to any of the plurality of servers, and a retransmitted data storage means for storing only the data where the confirmation reply is transmitted to the client about the data, out of the data transferred to the server, to be provided for the retransmission to the server.

COPYRIGHT: (C)2004,JPO



(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-94555

(P2004-94555A)

(43) 公開日 平成16年3月25日(2004.3.25)

(51) Int. Cl.⁷G06F 13/00
H04L 1/16
H04L 12/56

F I

G06F 13/00 353A
G06F 13/00 520C
G06F 13/00 520R
H04L 1/16
H04L 12/56 Z

テーマコード(参考)

5B089
5K014
5K030

審査請求 未請求 請求項の数 5 O L (全 25 頁)

(21) 出願番号 特願2002-254233 (P2002-254233)
(22) 出願日 平成14年8月30日(2002.8.30)(71) 出願人 000005223
富士通株式会社
神奈川県川崎市中原区上小田中4丁目1番
1号
(74) 代理人 100088244
弁理士 遠山 勉
(74) 代理人 100090516
弁理士 松倉 秀実
(72) 発明者 菊地 俊介
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内
(72) 発明者 小川 淳
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内
Fターム(参考) 5B089 GA11 HB02 KB05 JA11 KF02
最終頁に続く

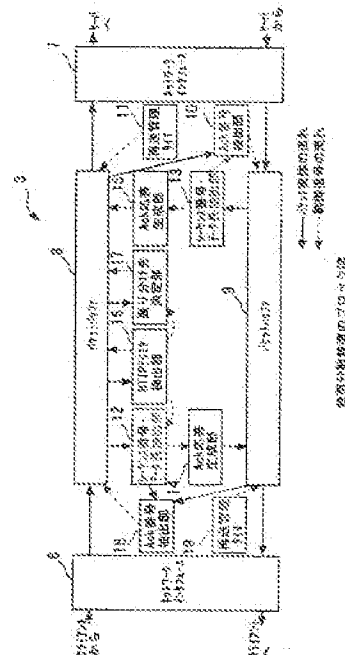
(54) 【発明の名称】 負荷分散装置

(57) 【要約】

【課題】クライアントから受信するリクエストに対する再送管理を省略することにより、送信バッファの使用時間を減少する。

【解決手段】クライアントから受信したデータの内容に応じて、クライアントへ当該データの確認応答を送信するか否かを判断する判断手段と、前記判断手段の判断結果に応じて、クライアントへ確認応答を送信する確認応答手段と、前記クライアントから受信したデータを複数のサーバのいずれかへ転送する転送手段と、前記サーバへ転送されたデータのうち、このデータについてクライアントへ確認応答が送信されたデータのみを、前記サーバへの再送に備えて記憶する再送データ記憶手段と、を備える。

【選択図】 図5



【特許請求の範囲】

【請求項1】

クライアントから受信したデータの内容に応じて、クライアントへ当該データの確認応答を送信するか否かを判断する判断手段と、
前記判断手段の判断結果に応じて、クライアントへ確認応答を送信する確認応答手段と、
前記クライアントから受信したデータを複数のサーバのいずれかへ転送する転送手段と、
前記サーバへ転送されたデータのうち、このデータについてクライアントへ確認応答が送信されたデータのみを、前記サーバへの再送に備えて記憶する再送データ記憶手段と、
を備える負荷分散装置。

【請求項2】

前記判断手段は、クライアントから受信したデータがデリミタを含まない場合にクライアントへ確認応答を送信すると判断し、前記データがデリミタを含む場合にクライアントへ確認応答を送信しないと判断する請求項1に記載の負荷分散装置。

【請求項3】

連続する複数のデータをサーバへ転送する場合に、最後にサーバに転送されたデータに対する確認応答をサーバから受信したときに、次のデータをサーバへ転送することを許可する確認応答判断手段をさらに備える請求項1又は2に記載の負荷分散装置。

【請求項4】

サーバから受信されたデータを、クライアントへ転送すべきか否かを判断する転送判断手段をさらに備える請求項1～3のいずれかに記載の負荷分散装置。

【請求項5】

前記転送判断手段は、前記サーバから受信されたデータが、確認応答のためのデータのみからなる場合に転送すべきでないと判断する請求項4に記載の負荷分散装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数のサーバを仮想的に一台のサーバのように見せて動作させる負荷分散装置に関する。

【0002】

【従来の技術】

任意の端末間におけるデータの送受信では、データ到着の保証を図るため、再送管理が行われている。再送管理を行う送信元端末は、送信バッファを有し、送信先端末へ送信したデータを、送信バッファに記憶しておく。そして、データの異常や消失が検出された場合に、送信元端末は送信先端末へ、送信バッファに記憶しておいたデータを再送する。送信先端末は、送信元端末から送信されたデータを正しく受信した場合に、確認応答データを送信元端末に送信する。送信元端末は、送信先端末から確認応答データを受信した場合に、送信バッファに記憶しておいたデータを消去し、送信バッファを解放する。このような再送管理を行うプロトコルの例として、TCP (Transport Control Protocol) がある。

【0003】

このような再送管理は、負荷分散装置でも行われる。負荷分散装置とは、複数のサーバに対して、クライアントからのリクエストを振り分ける装置である。負荷分散装置を用いることにより、クライアントからのリクエストが一定のサーバに集中することが回避される。

【0004】

このような負荷分散装置にとって、送信バッファを構成するメモリ領域は、有限のリソースである。一つのセッションにおける送信バッファの使用時間（保留時間）を減少させることにより、リソースの有効利用を図ることが可能となる。リソースの有効利用の結果として、例えばセッションの多重度の向上や、消費電力の減少などを図ることが可能となる。

10

20

30

40

50

【0005】

従来、TCPに対応する負荷分散装置において、再送管理を省略することにより送信バッファから早期にデータを消去し送信バッファの使用時間を減少させる技術がある。この技術では、負荷分散装置は、再送管理を省略する場合には、本来行う必要がある再送管理を、サーバ又はクライアントに任せる。この場合、負荷分散装置は、サーバからレスポンスを受信した際に、サーバに対しAck（確認応答）の送信を行わない。負荷分散装置は、サーバから受信したレスポンスをクライアントへ転送し、送信バッファに記憶しているレスポンスを消去する。

【0006】

図1は、従来技術における負荷分散装置P3が、再送管理を省略した場合の、クライアントP2、負荷分散装置P3、及びサーバP4の動作シーケンスを示す図である。図1を用いて、負荷分散装置P3が再送管理を省略する場合の、クライアントP2、負荷分散装置P3、及びサーバP4の動作について説明する。

【0007】

データが、サーバP4からクライアントP2に、正常に転送された場合（図1（a）参照）について説明する。まず、サーバP4は、負荷分散装置P3にレスポンスを送信する。負荷分散装置P3は、クライアントP2にレスポンスを転送する。クライアントP2は、レスポンスを受信し、負荷分散装置P3にAckを送信する。負荷分散装置P3は、サーバP4にAckを転送する。そして、サーバP4はAckを含むパケット（Ackパケット）を、負荷分散装置P3から受信する。このような動作によって、サーバP4は、データがクライアントP2に正常に受信されたことを確認する。

【0008】

次に、データがサーバP4－負荷分散装置P3間で消失した場合（図1（b）参照）について説明する。まず、サーバP4は、負荷分散装置P3にレスポンスを送信する。このとき、データが消失したと仮定する。この場合、負荷分散装置P3は何も受信せず、何の処理も行わない。そして、サーバP4は、タイムアウト秒が経過するまで待機しても、負荷分散装置P3からAckを受信できないため、レスポンスを再送する。

【0009】

次に、パケットが、負荷分散装置P3－クライアントP2間で消失した場合（図1（c）参照）について説明する。まず、サーバP4は、負荷分散装置P3にレスポンスを送信する。負荷分散装置P3は、受信したレスポンスをクライアントP2に転送する。このとき、データが消失したと仮定する。この場合、クライアントP2は何も受信せず、何の処理も行わない。そして、サーバP4はタイムアウト秒が経過するまで待機しても、負荷分散装置P3からAckを受信できないため、負荷分散装置P3へレスポンスを再送する。

【0010】

【発明が解決しようとする課題】

上記技術では、負荷分散装置は、サーバから受信するレスポンスについての再送管理を省略している。ここで、負荷分散装置は、クライアントから受信するリクエストについても、再送管理を省略することにより、送信バッファの使用時間をさらに減少させることができる。

【0011】

しかし、クライアントがデータ送信においてスロースタートメカニズムを用いている場合、負荷分散装置は、クライアントから受信するリクエストについての再送管理を省略することができない。

【0012】

図2は、スロースタートメカニズムを示す図である。図2を用いて、スロースタートメカニズムについて説明する。スロースタートメカニズムを用いた送信端末（クライアント）が受信端末（負荷分散装置）にデータを送信する場合、送信端末は、同時に送信するデータ数（同時送信データ数）を、1から指数関数的に増加させる（図2（a）参照）。このとき、送信端末は、受信端末から前のデータのAckを受信するまで、次のデータの送信

を待機する。

【0013】

クライアントP2が、スロースタートメカニズムを用いて、負荷分散装置P3に対し、データを送信すると仮定する。このとき、リクエストが単独のデータで構成されている場合、負荷分散装置P3は受信したデータについて、URLを抽出し、振り分け先を決定し、決定したサーバP4に対して転送処理を行うことが可能である。従って、最終的に、クライアントP2はサーバP4から、負荷分散装置P3を経由してレスポンスを受信できる。

【0014】

一方、リクエストが複数のデータに分割されて構成されている場合、クライアントP2は、リクエストを構成する1以上のデータを負荷分散装置P3に送信する毎に、そのデータ 10 に対するAckの受信を待機する状態となる(図2(b)参照)。図2(b)では、一つのリクエストが、データDcとデータDdとに分割されている。そして、クライアントP2は、データDcに対するAckを、負荷分散装置P3から受信するまで、データDdを送信しない。

【0015】

ここで、負荷分散装置P3がクライアントP2から受信するリクエストについての再送管理を省略している場合、クライアントP2は、送信したデータDcに対するAckを負荷分散装置P3から受信できない。このため、クライアントP2は、負荷分散装置P3へ、データDcの再送を繰り返す。再送が繰り返されても、負荷分散装置P3は、クライアントP2にAckを送信しない。この結果、負荷分散装置P3は、データDdを受信できず 20 、リクエストを完全に受信することができない。このため、負荷分散装置P3は、リクエストについて、URLの抽出や、振り分け先の決定などを実行できない。従って、負荷分散装置P3は、クライアントP2から受信したリクエストを、サーバP4に転送できず、全体としてデータの送受信が滞ってしまう。

【0016】

負荷分散装置P3は、上記したようなリクエストを受信できなくなることを避けるため、クライアントP2から受信するリクエストについては、そのリクエストの内容に関わらず、クライアントP2にAckを送信する。この場合、負荷分散装置P3は、クライアントP2から受信したリクエストについて再送管理を行う必要がある。即ち、負荷分散装置P3は、サーバP4から、リクエストに対するAckを受信するまで、クライアントP2から 30 受信したリクエストを送信バッファに記憶しておく必要がある。

【0017】

図3は、従来技術の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。図3を用いて、負荷分散装置P3がクライアントP2から受信したリクエストに対してAckを送信するにも関わらずこのリクエストを送信バッファに記憶しない場合に発生する問題について説明する。負荷分散装置P3からサーバP4に転送されたリクエストが、正常にサーバP4で受信された場合(図3(a)参照)には、問題は生じない。

【0018】

一方、負荷分散装置P3からサーバP4に転送されたリクエストがサーバP4に受信される前に消失した場合(図3(b)参照)には、次のような問題が発生する。即ち、負荷分散 40 装置P3は、送信バッファにリクエストを記憶していないため、サーバP4にリクエストを再送できない。一方、クライアントP2は、負荷分散装置P3からAckを受信した時点で、リクエストを送信バッファから消去する。このため、クライアントP2は、負荷分散装置P3にリクエストを再送できない。

【0019】

従って、負荷分散装置P3は、リクエストの消失に備え、サーバP4からAckを受信するまで、リクエストを送信バッファから消去せずに記憶しておかなければならない。このように、従来技術における負荷分散装置P3は、クライアントP2から受信するリクエストに対する再送管理を省略できない。このため、従来技術における負荷分散装置P3は、送信バッファの使用時間を減少できない。

【0020】

本発明は、このような問題を解決し、クライアントから受信するリクエストに対する再送管理を省略することにより、送信バッファの使用時間を減少することが可能な負荷分散装置を提供することを目的とする。

【0021】

【課題を解決するための手段】

上記問題を解決するため、本発明は以下のような構成をとる。本発明の第一の態様は、負荷分散装置であって、クライアントから受信したデータの内容に応じて、クライアントへ当該データの確認応答を送信するか否かを判断する判断手段と、前記判断手段の判断結果に応じて、クライアントへ確認応答を送信する確認応答手段と、前記クライアントから受信したデータを複数のサーバのいずれかへ転送する転送手段と、前記サーバへ転送されたデータのうち、このデータについてクライアントへ確認応答が送信されたデータのみを、前記サーバへの再送に備えて記憶する再送データ記憶手段と、を備える。

【0022】

本発明の第一の態様によれば、判断手段は、クライアントから受信したデータの内容を解析し、このデータの内容に応じてクライアントへ確認応答を送信するか否かを判断する。この確認応答は、クライアントから受信したデータに対する確認応答である。判断手段が上記判断を実行すると、確認応答手段は、この判断結果に基づいて、クライアントへ確認応答を送信する。転送手段はクライアントから受信したデータを複数のサーバのいずれかへ転送する。再送データ記憶手段は、サーバへ転送されたデータを、このサーバへのデータ再送に備えて記憶する。

【0023】

このため、再送データ記憶手段は、確認応答がクライアントに送信されないデータを記憶しない。従って、再送データ記憶手段は、サーバに転送されるデータ全てを記憶するわけではなく、使用される記憶領域が削減される。

【0024】

また、本発明の第一の態様における前記判断手段は、クライアントから受信したデータがデリミタを含まない場合にクライアントへ確認応答を送信すると判断し、前記データがデリミタを含む場合にクライアントへ確認応答を送信しないと判断するように構成されても良い。

【0025】

このように構成された本発明の第一の態様によれば、判断手段は、クライアントから受信したデータがデリミタを含むか否かに応じて、クライアントへこのデータの確認応答を送信するか否かを判断する。そして、確認応答手段は、デリミタを含まないデータが受信された場合、このデータに対する確認応答をクライアントへ送信する。

【0026】

このため、クライアントがスロースタートメカニズムを採用している場合であっても、負荷分散装置はクライアントからデリミタを含まないデータの次のデータを受信することが可能となる。

【0027】

また、本発明の第一の態様は、連続する複数のデータをサーバへ転送する場合に、最後にサーバに転送されたデータに対する確認応答をサーバから受信したときに、次のデータをサーバへ転送することを許可する確認応答判断手段をさらに備えるように構成されても良い。

【0028】

また、本発明の第一の態様は、サーバから受信されたデータを、クライアントへ転送すべきか否かを判断する転送判断手段をさらに備えるように構成されても良い。

【0029】

また、本発明の第一の態様における前記転送判断手段は、前記サーバから受信されたデータが、確認応答のためのデータのみからなる場合に転送すべきでないと判断するように構

成されても良い。

【0030】

【発明の実施の形態】

次に、図を用いて本発明の実施形態における負荷分散装置を用いた負荷分散システムについて説明する。なお、本実施形態の説明は例示であり、本発明は以下の説明する構成に限定されない。

【0031】

【第一実施形態】

＜システム構成＞

図4は、負荷分散システム1の概略を示す図である。第一実施形態における負荷分散装置3を用いた負荷分散システム1は、複数のクライアント2、負荷分散装置3、複数のサーバ4、及び複数のクライアント2と負荷分散装置3とを通信可能に接続するネットワーク5を用いて構成される。本実施形態における負荷分散装置3は、例として、TCP/IP (Transport Control Protocol / Internet Protocol) を用いてHTTPを取り扱う。これに従い、クライアント2とサーバ4とは、少なくとも、TCP/IPを用いてHTTPを取り扱う。負荷分散装置3は、確認応答を行うプロトコルを扱う装置であれば良く、TCP/IPを取り扱う装置に限定されない。以下、負荷分散システム1の各構成について説明する。

【0032】

＜＜クライアント＞＞

クライアント2は、パーソナルコンピュータやワークステーション等の情報処理装置を用いて構成される。クライアント2は、ネットワークカード等の通信装置を用いて、ネットワーク5を介して負荷分散装置3と通信可能に構成される。クライアント2は、ブラウザソフトがインストールされており、負荷分散装置3に対し、HTTPリクエストを含むパケットを送信する。そして、クライアント2は、負荷分散装置3から、このHTTPリクエストに対するHTTPレスポンスを含むパケットを受信する。クライアント2は、受信したHTTPレスポンスの内容を、ブラウザソフトを用いて表示する。また、クライアント2は、スロースタートメカニズムを用いてパケットを送信する。

【0033】

＜＜負荷分散装置＞＞

負荷分散装置3は、パーソナルコンピュータやワークステーション等の情報処理装置や、負荷分散処理を行うことを前提とした専用ハードウェア等によって構成される。図5は、負荷分散装置3のブロック図である。図5を用いて負荷分散装置3について説明する。負荷分散装置3は、ハードウェア的には、バスを介して接続されたCPU、主記憶(RAM)、補助記憶装置(ハードディスク)、通信制御装置(ネットワークカード等)等を備えている。あるいは、専用のパケット処理プロセッサ、パケット処理のための記憶装置(RAM)、通信制御装置(ネットワークカード等)等から構成される場合もある。負荷分散装置3は、補助記憶装置に記憶された各種のプログラム(OS、アプリケーション等)が主記憶にロードされCPUにより実行されることによって、あるいはパケット処理プロセッサがあらかじめ組み込まれた処理に従って動作することによって、ネットワークインタフェース6、7、パケットバッファ8、9、Ack番号抽出部10、18、再送管理タイマ11、19、シーケンス番号・データ長抽出部12、13、Ack応答生成部14、15、HTTPデリミタ検出部16、及び振り分け先決定部17等を含む装置として機能する。

【0034】

ネットワークインタフェース6、7は、ネットワークカード等を用いて構成される。ネットワークインタフェース6は、クライアント2から送信されたパケットを、ネットワーク5を介して受信し、パケットバッファ8に書き込む。また、ネットワークインタフェース6は、パケットバッファ9から受信するパケットを、ネットワーク5を介してクライアント2に送信する。ネットワークインタフェース7は、サーバ4から送信されたパケットを

受信し、パケットバッファ9に書き込む。また、ネットワークインタフェース7は、パケットバッファ8から受信するパケットを、サーバ4に送信する。

【0035】

パケットバッファ8、9は、CPUやRAM等、あるいは専用パケット処理プロセッサ等を用いて構成される。パケットバッファ8は、再送管理タイマ11から再送信信号を受信すると、この再送信信号に含まれる識別子に対応するパケットを、ネットワークインタフェース7へ送信する。パケットバッファ8は、クライアント2から送信されたパケット、あるいは負荷分散装置3によって生成されたパケットをバッファリングする。そして、パケットバッファ8は、バッファリングしているパケットをネットワークインタフェース7へ送信する。また、パケットバッファ8は、Ack応答生成部15によってパケットをバッファリングされると、このパケットをネットワークインタフェース7へ送信する。また、パケットバッファ8は、振り分け先決定部17から消去信号を受信すると、この消去信号に含まれる識別子に対応するパケットが転送されたことを確認し、このパケットを消去する。

【0036】

パケットバッファ9は、再送管理タイマ19から再送信信号を受信すると、この再送信信号に含まれる識別子に対応するパケットを、ネットワークインタフェース6へ送信する。パケットバッファ9は、サーバ4から送信されたパケット、あるいは負荷分散装置3によって生成されたパケットをバッファリングする。そして、パケットバッファ9は、バッファリングしているパケットをネットワークインタフェース6へ送信する。また、パケットバッファ9は、Ack応答生成部14によってパケットをバッファリングされると、このパケットをネットワークインタフェース6へ送信する。

【0037】

Ack番号抽出部10、18は、CPUやRAM等、あるいは専用パケット処理プロセッサ等を用いて構成される。Ack番号抽出部10は、パケットバッファ8にバッファリングされたパケットのAck番号を抽出する。また、Ack番号抽出部10は、シーケンス番号・データ長抽出部13から、シーケンス番号及びデータ長を受信する。即ち、Ack番号抽出部10は、パケットバッファ9にバッファリングされたパケットのシーケンス番号及びデータ長を受信する。そして、Ack番号抽出部10は、Ack番号の値と、シーケンス番号及びデータ長の和とを比較する。Ack番号抽出部10は、Ack番号の方が大きい値又は同じ値である場合、パケットバッファ9に、このシーケンス番号及びデータ長の値を含むパケットの識別子を含む消去信号を送信する。消去信号は、消去すべきパケットの識別子を含む。

【0038】

Ack番号抽出部18は、パケットバッファ9にバッファリングされたパケットのAck番号を抽出する。また、Ack番号抽出部18は、シーケンス番号・データ長抽出部12から、シーケンス番号及びデータ長を受信する。即ち、Ack番号抽出部18は、パケットバッファ8にバッファリングされたパケットのシーケンス番号及びデータ長を受信する。そして、Ack番号抽出部18は、Ack番号の値と、シーケンス番号とデータ長とを加算した値とを比較する。Ack番号抽出部18は、Ack番号の方が大きい値又は同じ値である場合、パケットバッファ8に消去信号を送信する。消去信号は、消去すべきパケットの識別子を含む。

【0039】

再送管理タイマ11、19は、クロックやCPU等、あるいは専用パケット処理プロセッサ等を用いて構成される。再送管理タイマ11は、パケットバッファ8にバッファリングされたパケットが、ネットワークインタフェース7によって送信されてからの経過時間を計時する。再送管理タイマ11は、送信されてから一定時間以上経過した（タイムアウトの）パケットがある場合、このパケットの識別子を含む再送信信号を、パケットバッファ8へ送信する。

【0040】

再送管理タイマ19は、パケットバッファ9にバッファリングされたパケットが、ネットワークインタフェース6によって送信されてからの経過時間を計時する。再送管理タイマ19は、送信されてから一定時間以上経過したパケットがある場合、このパケットの識別子を含む再送信信号を、パケットバッファ9へ送信する。

【0041】

シーケンス番号・データ長抽出部12、13は、CPUやRAM等、あるいは専用パケット処理プロセッサ等を用いて構成される。シーケンス番号・データ長抽出部12は、パケットバッファ8にバッファリングされているパケットのシーケンス番号及びデータ長を抽出する。そして、シーケンス番号・データ長抽出部12は、シーケンス番号及びデータ長を、Ack応答生成部14及びAck番号抽出部18に送信する。

10

【0042】

シーケンス番号・データ長抽出部13は、パケットバッファ9にバッファリングされているパケットのシーケンス番号及びデータ長を抽出する。そして、シーケンス番号・データ長抽出部13は、シーケンス番号及びデータ長を、Ack応答生成部15及びAck番号抽出部10に送信する。

【0043】

Ack応答生成部14、15は、CPUやRAM等、あるいは専用パケット処理プロセッサ等を用いて構成される。Ack応答生成部14は、シーケンス番号・データ長抽出部12から受信した二つの値の和をAck番号として持つパケットを生成する。そして、Ack応答生成部14は、生成したパケットをパケットバッファ9にバッファリングする。

20

【0044】

Ack応答生成部15は、シーケンス番号・データ長抽出部13から受信した二つの値の和をAck番号として持つパケットを生成する。そして、Ack応答生成部15は、生成したパケットをパケットバッファ8にバッファリングする。

【0045】

HTTPデリミタ検出部16は、CPUやRAM等、あるいは専用パケット処理プロセッサ等を用いて構成される。HTTPデリミタ検出部16は、パケットバッファ8にバッファリングされるHTTPリクエストを含むパケット（リクエストパケット）のペイロードに、HTTPのデリミタ（Delimiter：例えば空行）が存在するか否かを判断する。HTTPデリミタ検出部16は、デリミタが存在する場合、振り分け先決定部17へ、転送信号を送信する。この転送信号は、HTTPデリミタ検出部16がデリミタを検出したリクエストパケットの識別子を含む。

30

【0046】

また、HTTPデリミタ検出部16は、リクエストパケットにデリミタが存在しない場合、シーケンス番号・データ長抽出部12へ指示を出し、クライアント2に対し、Ackを含むパケット（Ackパケット）を送信する処理を実行させる。

【0047】

振り分け先決定部17は、CPUやRAM等、あるいは専用パケット処理プロセッサ等を用いて構成される。振り分け先決定部17は、HTTPデリミタ検出部16から転送信号を受信すると、転送信号に含まれる識別子に対応するリクエストパケットに含まれるURLを抽出し、このURLが示すコンテンツを保有する最適なサーバ4を選択する。その上でそのリクエストパケットを、選択したサーバ4へ転送する。即ち、振り分け先決定部17は、転送対象となるリクエストパケットの送信先アドレスや送信元アドレス等の情報を書き換えることにより、このリクエストパケットをサーバ4へ転送する。また、振り分け先決定部17は、転送対象となるリクエストパケットのうち、デリミタを有するリクエストパケットの識別子を含む消去信号をパケットバッファ8へ送信する。

40

【0048】

なお、Ack番号抽出部10、18、シーケンス番号・データ長抽出部12、13、HTTPデリミタ検出部16、振り分け先決定部17、及びAck応答生成部14、15は、各々独立に動作するのではなく、パケットの到着に伴い順に動作する。また、パケット受

50

信、送信に伴う上記記載以外に必要な処理については、従来行われている処理であるため、説明を省略する。

【0049】

〈サーバ〉

図4に戻って、サーバ4は、パーソナルコンピュータやワークステーション等の情報処理装置を用いて構成される。サーバ4は、ネットワークカード等の通信装置を用いて、負荷分散装置3と通信可能に構成される。サーバ4では、HTTPD (HTTP Daemon) が実行される。サーバ4は、負荷分散装置3から受信するリクエストパケットに対し、HTTPレスポンスを含むパケット (レスポンスパケット) を返信する。

【0050】

〈動作例〉

図6～8は、負荷分散装置3の動作例を示すフローチャートである。図5、図6～8を用いて、負荷分散装置3の動作例について説明する。ここでは、負荷分散装置3は、クライアント2からリクエストパケットを受信し、サーバ4からレスポンスパケットを受信するものとして説明する。

【0051】

負荷分散装置3のネットワークインタフェース6は、クライアント2からリクエストパケットを受信する (図6参照: S01)。パケットバッファ8は、受信したリクエストパケットをバッファリングする (S02)。Ack番号抽出部10は、このリクエストパケットからAck番号を抽出する (S03)。また、シーケンス番号・データ長抽出部13は、パケットバッファ9にバッファリングされたパケットのシーケンス番号及びデータ長を抽出し、Ack番号抽出部10へ送信する (S04)。Ack番号抽出部10は、抽出したAck番号と、シーケンス番号及びデータ長の和とを比較する (S05)。Ack番号よりもシーケンス番号及びデータ長の和が小さな値となるパケットがパケットバッファ9にバッファリングされている場合 (S05-YES)、パケットバッファ9は、そのパケットを消去する (S06)。

【0052】

次に、HTTPデリミタ検出部16は、パケットバッファ8にバッファリングされたリクエストパケットがデリミタを含むか否かを判断する (S07)。HTTPデリミタ検出部16がデリミタを検出しない場合 (S07-NO)、シーケンス番号・データ長抽出部12は、Ack応答生成部14へ、抽出したシーケンス番号及びデータ長を送信する。Ack番号応答生成部14は、受信したシーケンス番号及びデータ長を加算する。そして、Ack番号応答生成部14は、この値をAck番号として持つパケットを生成し、パケットバッファ9にこのパケットをバッファリングする (S08)。パケットバッファ9は、バッファリングされたパケットを、ネットワークインタフェース6を介して、クライアント2へ送信する (S09)。この後、負荷分散装置3は、新たなリクエストパケットをクライアント2から受信すると、改めてS01～S07の処理を実行する。

【0053】

一方、HTTPデリミタ検出部16は、デリミタを検出した場合 (S07-YES)、振り分け先決定部17へ転送信号を送信する。振り分け先決定部17は、パケットバッファ8からリクエストパケットのペイロード部を読み込み、そこに示されているURLをコンテンツとして持つ適当なサーバ4を決定する (S10)。振り分け先決定部17は、決定されたサーバ4に対するTCPコネクションがオープンされると、サーバ4に対し、リクエストパケット全てを転送する (S11)。このとき、振り分け先決定部17は、転送対象となるリクエストパケットのうち、デリミタを有するリクエストパケットの識別子を含む消去信号をパケットバッファ8へ送信する。そして、パケットバッファ8はこのリクエストパケットを消去する。

【0054】

再送管理タイマ11は、タイムアウトのリクエストパケットがある場合 (図7参照: S12-YES)、このリクエストパケットの識別子を含む再送信号を、パケットバッファ8

10

20

30

40

50

へ送信する。そして、バケットバッファ8は、この再送信号に含まれる識別子が示すリクエストバケットを、サーバ4へ再送信する(S13)。

【0055】

一方、タイムアウトのリクエストバケットが無い場合(S12-No)、ネットワークインタフェース7がサーバ4からバケットを受信したか否かを判断する(S14)。ネットワークインタフェース7がバケットを受信していない場合(S14-No)、再びタイムアウトの判断に戻る。一方、ネットワークインタフェース7がサーバ4からバケットを受信した場合(S14-Yes)、バケットバッファ9は、このバケットをバッファリングする(S15)。

【0056】

バケットバッファ9がバケットをバッファリングした後、Ack番号抽出部18は、バッファリングされたバケットのAck番号を抽出する(図8参照:S16)。また、シーケンス番号・データ長抽出部12は、バケットバッファ8にバッファリングされたバケットのシーケンス番号及びデータ長を抽出し、Ack番号抽出部18へ送信する(S17)。Ack番号抽出部18は、シーケンス番号及びデータ長の和が、抽出されたAck番号と同じ又は小さい値であるバケットの存在を判断する(S18)。このようなバケットがバケットバッファ8にバッファリングされている場合(S18-Yes)、バケットバッファ8は、そのバケットを消去する(S19)。

【0057】

バケットバッファ9は、バッファリングしたバケット(サーバ4から受信されたバケット)を、クライアント2に転送する(S20)。そして、バケットバッファ9は、転送したバケットを消去する(S21)。

【0058】

〈動作シーケンス〉

次に、第一実施形態における負荷分散システム1の動作シーケンスについて説明する。図9(a)、(b)は、HTTPリクエストが一つのリクエストバケットで構成される場合の、負荷分散システム1の動作シーケンスを示す図である。第一に、図9(a)を用いて、HTTPリクエストが一つのリクエストバケットで構成され、リクエストバケットが消失しない場合の動作シーケンスについて説明する。

【0059】

まず、クライアント2と負荷分散装置3とは、TCPのコネクションをオープンする(オープンフェーズ:図9(a)参照)。クライアント2は、コネクションをオープンした後、データ転送フェーズの処理として、負荷分散装置3にリクエストバケット(HTTPリクエストを含むバケット)を送信する。

【0060】

負荷分散装置3は、クライアント2からリクエストバケットを受信する。負荷分散装置3は、受信したリクエストバケットがデリミタを含むか否かを判断する。この場合、リクエストバケットは一つのバケットで構成されるため、リクエストバケットはデリミタを含む。このため、負荷分散装置3は、リクエストバケットに示されているURLをコンテンツとして持つ適当なサーバ4を決定する。そして、負荷分散装置3は、決定されたサーバ4に対するTCPコネクションをオープンし、サーバ4に対し、リクエストバケットを転送する。このとき、負荷分散装置3は、このリクエストバケットを、バケットバッファ8から消去する。

【0061】

サーバ4は、負荷分散装置3からリクエストバケットを受信する。サーバ4は、受信したリクエストバケットに対するレスポンスバケット(HTTPレスポンスを含むバケット)を生成する。そして、サーバ4は、負荷分散装置3に対し、レスポンスバケットを送信する。

【0062】

負荷分散装置3は、レスポンスバケットを受信する。負荷分散装置3は、受信したレスポ

10

20

30

40

50

ンスパケットを、クライアント2に転送する。このとき、負荷分散装置3は、このレスポンスパケットをパケットバッファ9から消去する。そして、クライアント2は、クローズフェーズの処理を実行する。

【0063】

第二に、図9(b)を用いて、HTTPリクエストが一つのリクエストパケットで構成され、リクエストパケットが消失する場合の動作シーケンスについて説明する。ただし、負荷分散システム1の動作シーケンスにおけるオープンフェーズとクローズフェーズとは、どのような場合であっても同じであるため、以後、オープンフェーズとクローズフェーズについての説明を省略する。

【0064】

まず、クライアント2はリクエストパケットを負荷分散装置3へ送信する。負荷分散装置3は、リクエストパケットのペイロードを解析し、振り分け先サーバ4を決定し、サーバ4に対してTCPコネクションをオープンする。負荷分散装置3はリクエストパケットを、決定されたサーバ4に転送する。このとき、転送されたリクエストパケットが消失したとする。この場合、サーバ4は、負荷分散装置3から、転送されたリクエストパケットを受信しない。従って、サーバ4は、リクエストパケットに対するAckパケットを、負荷分散装置3へ送信しない。このため、この場合、クライアント2は、リクエストパケットに対するAckパケットを、負荷分散装置3から受信できない。クライアント2は、タイムアウトとなると(タイムアウト秒待つと)、リクエストパケットを、負荷分散装置3へ再送する。そして、負荷分散装置3は、再送されたリクエストパケットを受信し、サーバ4へ転送する。クライアント2は、送信したリクエストパケットに対するAckを受信するまでこの動作を繰り返す。そして、最終的に、クライアント2は、負荷分散装置3から、レスポンスパケットを受信する。

【0065】

図10、11(a)、11(b)は、HTTPリクエストが複数のリクエストパケットで構成される場合の、負荷分散システム1の動作シーケンスを示す図である。第三に、図10を用いて、HTTPリクエストが複数のリクエストパケットで構成され、リクエストパケットが消失しない場合の動作シーケンスについて説明する。

【0066】

ここではデリミタを含まないリクエストパケットをリクエストパケットa(HTTPリクエストaを含むリクエストパケット)、デリミタを含むリクエストパケットをリクエストパケットb(HTTPリクエストbを含むリクエストパケット)とする。即ち、一つのHTTPリクエストが、HTTPリクエストaとHTTPリクエストbとに分かれて構成される。HTTPリクエストが、三つ以上のリクエストパケットで構成されている場合、最後のリクエストパケットをリクエストパケットb、その直前までのリクエストパケットをリクエストパケットaとみなせばよい。

【0067】

クライアント2は、リクエストパケットaを、負荷分散装置3へ送信する。負荷分散装置3はリクエストパケットaのペイロードを解析し、リクエストパケットaにデリミタが含まれていないと判断する。このため、負荷分散装置3は、クライアント2に対しAckを含むパケット(Ackパケット)を送信する。

【0068】

クライアント2は、負荷分散装置3から、Ackパケットを受信する。そして、クライアント2は、負荷分散装置3へ、リクエストパケットbを送信する。負荷分散装置3は、リクエストパケットbのペイロードを解析し、デリミタが含まれていると判断する。このため、負荷分散装置3は、先に受信したリクエストパケットaとリクエストパケットbとを用いて、振り分け先サーバ4を決定する。そして、負荷分散装置3は、決定したサーバ4に対してTCPコネクションをオープンし、リクエストパケットa、bをサーバ4に転送する。

【0069】

10

20

30

40

50

サーバ4は、リクエストパケットa、bを、負荷分散装置3から受信する。サーバ4は、受信したリクエストパケットに対するレスポンスパケットを生成する。そして、サーバ4は、生成したレスポンスパケットを、負荷分散装置3へ送信する。

【0070】

負荷分散装置3は、レスポンスパケットに含まれるAck番号を用いて、リクエストパケットaに対する受信確認をする。負荷分散装置3は、レスポンスパケットをクライアント2に転送する。

【0071】

クライアント2は、レスポンスパケットを受信する。そして、クライアント2は、受信したレスポンスパケットにより、受信確認をする。

10

【0072】

第四に、図11(a)を用いて、HTTPリクエストが複数のリクエストパケットで構成され、リクエストパケットaが消失する場合の動作シーケンスについて説明する。ただし、負荷分散装置3がリクエストパケットa、bをサーバ4に転送するまでの動作シーケンスは、第三の場合と同じであるため、その説明を省略する。

【0073】

負荷分散装置3がリクエストパケットa、bをサーバ4に転送した後に、リクエストパケットaが消失したものとする。この場合、サーバ4は、リクエストパケットbを受信するが、リクエストパケットaを受信しない。このため、サーバ4は、リクエストパケットaに対するAckを含むパケットを送信しない。

20

【0074】

負荷分散装置3は、サーバ4から、リクエストパケットaに対するAckを含むパケットを受信できないため、タイムアウト秒経過後、リクエストパケットaをサーバ4に再送する。負荷分散装置3は、サーバ4から、リクエストパケットaに対するAckを含むパケットを受信するまでこの処理を繰り返す。最終的に、負荷分散装置3は、サーバ4からレスポンスパケットを受信し、受信したレスポンスパケットをクライアント2へ転送する。このとき、このレスポンスパケットは、リクエストパケットaに対するAckを含む。このため、負荷分散装置3は、サーバ4から、リクエストパケットaに対するAckを含むパケットを受信したこととなる。

【0075】

第五に、図11(b)を用いて、HTTPリクエストが複数のリクエストパケットで構成され、リクエストパケットbが消失する場合の動作シーケンスについて説明する。ただし、負荷分散装置3がリクエストパケットa、bをサーバ4に転送するまでの動作シーケンスは、第三の場合と同じであるため省略する。

30

【0076】

負荷分散装置3がリクエストパケットa、bをサーバ4に転送した後に、リクエストパケットbが消失したものとする。この場合、サーバ4は、リクエストパケットaのみを受信する。このため、サーバ4はリクエストパケットaのみに対するAckを含むパケットを、負荷分散装置3へ送信する。

【0077】

負荷分散装置3は、リクエストパケットaに対するAckを含むパケットを受信する。その後、負荷分散装置3は、このAckを含むパケットをクライアント2へ転送する。

40

【0078】

クライアント2は、リクエストパケットaに対するAckを含むパケットを受信する。しかし、クライアント2は、リクエストパケットaに対するAckを含むパケットは既に受信済みであるため、何も実行しない。即ち、クライアント2は、既に負荷分散装置3から、リクエストパケットaに対するAckを含むパケットを受信しているため、何の処理も実行しない。ただし、クライアント2は、リクエストパケットbに対するAckを含むパケットを受信できない。このため、クライアント2は、タイムアウト秒経過後、リクエストパケットbを、負荷分散装置3へ再送する。

50

【0079】

クライアント2は、負荷分散装置3からレスポンスパケットを受信するまでこの処理を繰り返す。このとき、レスポンスパケットは、リクエストパケットbに対するAckを含む。このため、クライアント2は、負荷分散装置3からAckパケットを受信したこととなる。

【0080】

〈作用〉

本発明の第一実施形態における負荷分散装置3は、HTTPリクエストが一つのリクエストパケットにより構成される場合、リクエストパケットに対するAckをクライアント2に送信しない。従って、負荷分散装置3は、リクエストパケットについて再送管理を実行する必要が無い。このため、このリクエストパケットを、パケットバッファ8から早期に消去することが可能となる。即ち、負荷分散装置3は、パケットバッファ8の記憶領域を、早期に解放することが可能となる。

【0081】

図12(a)は、HTTPリクエストが一つのリクエストパケットにより構成される場合に、パケットバッファ8の記憶領域が解放されるタイミングを示す図である。従来の負荷分散装置は、リクエストパケットに対するAckをサーバから受信した時点（レスポンスパケットを受信した時点）で、リクエストパケットをパケットバッファ8から消去する。一方、負荷分散装置3は、リクエストパケットをサーバ4に転送した時点で、リクエストパケットをパケットバッファ8から消去する。このタイミングの差分を、図12(a)に太線として示す。このように、本発明による負荷分散装置3によれば、パケットバッファ8の記憶領域を早期に解放することが可能となる。

【0082】

また、HTTPリクエストが複数のリクエストパケットにより構成される場合、負荷分散装置3は、最後のリクエストパケット（デリミタを含むリクエストパケット）に対するAckをクライアント2に送信しない。このため、負荷分散装置3は、最後のリクエストパケットについて再送管理を実行する必要が無い。従って、最後のリクエストパケットをパケットバッファ8から早期に消去することが可能となる。即ち、負荷分散装置3は、パケットバッファ8の記憶領域を早期に解放することが可能となる。

【0083】

図12(b)は、HTTPリクエストが複数のリクエストパケットにより構成される場合に、パケットバッファ8の記憶領域が解放されるタイミングを示す図である。従来の負荷分散装置は、最後のリクエストパケットに対するAckをサーバから受信した時点（レスポンスパケットを受信した時点）で、パケットバッファ8から最後のリクエストパケットを消去する。しかし、負荷分散装置3は、最後のリクエストパケットをサーバ4に転送した時点で、パケットバッファ8から最後のリクエストパケットを消去する。このタイミングの差分を、図12(b)に太線として示す。

【0084】

このように、第一実施形態の負荷分散装置3は、パケットバッファ8の記憶領域を早期に開放する。このため、負荷分散装置3におけるセッションの多重率の向上、それに伴う負荷分散装置3におけるレイテンシの向上、発熱量、消費電力、実装面積の縮小などが実現される。

【0085】

〈変形例〉

本発明の第一実施形態による負荷分散装置3は、パケットバッファ8、9がそれぞれ別に構成されているが、一つの記憶装置を用いて構成されても良い。

【0086】

〔第二実施形態〕

〈システム構成〉

第二実施形態における負荷分散装置3aを用いた負荷分散システム1aについて説明する

。負荷分散システム 1 a のシステム構成は、基本的に、第一実施形態における負荷分散装置 3 を用いた負荷分散システム 1 と同じである。即ち、負荷分散システム 1 a は、複数のクライアント 2、負荷分散装置 3 a、複数のサーバ 4、及び複数のクライアント 2 と負荷分散装置 3 a とを通信可能に接続するネットワーク 5 とを用いて構成される。クライアント 2 とサーバ 4 とは、上記した第一実施形態においてすでに説明したため、以下では負荷分散装置 3 a の構成について説明する。

【0087】

〈〈負荷分散装置〉〉

図 13 は、本発明による第二実施形態の負荷分散装置 3 a のブロック図である。

図 13 を用いて、負荷分散装置 3 a の構成について説明する。ただし、負荷分散装置 3 a 10 について、負荷分散装置 3 と異なる構成についてのみ説明する。負荷分散装置 3 a は、パケット送信許可部 20 をさらに備える。

【0088】

パケット送信許可部 20 は、CPU や RAM 等、あるいは専用パケット処理プロセッサ等を用いて構成される。パケット送信許可部 20 は、サーバ 4 に転送されたリクエストパケットのシーケンス番号とデータ長とを記憶する。また、パケット送信許可部 20 は、サーバ 4 から受信したパケットの Ack 番号と、シーケンス番号及びデータ長の和とを比較する。パケット送信許可部 20 は、二つの値が一致した場合、パケットバッファ 8 に対し、後続のリクエストパケットをサーバ 4 に転送することを許可する。

【0089】

20

〈動作例〉

図 14 は、第二実施形態における負荷分散装置 3 a の動作例のうち、負荷分散装置 3 と異なる部分のみを示したフローチャートである。図 14 を用いて、負荷分散装置 3 a の動作例について説明する。ただし、負荷分散装置 3 の動作と異なる動作についてのみ説明する。

【0090】

振り分け先決定部 17 は、振り分け先となるサーバ 4 を決定すると (S10)、決定されたサーバ 4 に対し、最初のリクエストパケットを一つ転送する (S22)。このとき、パケット送信許可部 20 は、転送されたリクエストパケットにおけるシーケンス番号及びデータ長を記憶する。

30

【0091】

パケットバッファ 9 に Ack を含むパケットがバッファリングされると、パケット送信許可部 20 は、このパケットの Ack 番号と、自身が記憶するシーケンス番号及びデータ長の和とを比較する。この値が一致した場合、パケット送信許可部 20 は、転送されたリクエストパケットに対する Ack を含むパケットを受信したと判断する (S23)。

【0092】

パケット送信許可部 20 が、転送されたリクエストパケットに対する Ack を含むパケットを受信していないと判断した場合、又はパケットバッファ 9 に Ack を含むパケットがバッファリングされていない場合 (S23-No)、再送管理タイマ 11 は、転送されたリクエストパケットについてタイムアウトが否かを判断する (S24)。タイムアウトである場合 (S24-Yes)、パケットバッファ 8 は、タイムアウトと判断されたリクエストパケットを、サーバ 4 へ再送する (S22)。一方、タイムアウトでない場合 (S24-No)、負荷分散装置 3 a は、転送されたリクエストパケットに対する Ack が受信されるまで待機する。

40

【0093】

パケット送信許可部 20 が、転送されたリクエストパケットに対する Ack を含むパケットを受信したと判断した場合 (S23-Yes)、パケット送信許可部 20 は、全てのリクエストパケットが転送されたか否かを判断する (S25)。転送されていないリクエストパケットがある場合 (S25-No)、パケット送信許可部 20 は、後続のリクエストパケットが転送されることを許可する。一方、全てのリクエストパケットが転送された場 50

合（S25-Yes）、S12以降の処理が実行される（図6参照）。

【0094】

＜動作シーケンス＞

図15は、第二実施形態における負荷分散装置3aを用いた負荷分散システム1aの動作シーケンスを示す図である。図15を用いて、負荷分散システム1aの動作シーケンスについて説明する。ただし、負荷分散システム1の動作シーケンスと異なる動作シーケンスについてのみ説明する。

【0095】

負荷分散装置3aは、リクエストパケットの転送先となるサーバ4を決定した後、決定したサーバ4に対し、リクエストパケットaを転送する。サーバ4は、負荷分散装置3aからリクエストパケットaを受信する。そして、サーバ4は、受信したリクエストパケットaに対するAckを含むパケットを、負荷分散装置3aに送信する。

【0096】

負荷分散装置3aは、サーバ4から、リクエストパケットaに対するAckを含むパケットを受信すると、このパケットをクライアント2へ転送する。そして、負荷分散装置3aは、後続のリクエストパケット、即ちリクエストパケットbをサーバ4へ転送する。

【0097】

＜作用＞

図16は、従来技術における負荷分散装置P3と本発明の第二実施形態における負荷分散装置3aとの動作の差異を示す図である。図16を用いて、本発明の第二実施形態における負荷分散装置3aの作用について説明する。従来技術における負荷分散装置P3は、複数のリクエストパケットで構成されるHTTPリクエストをサーバP4へ転送する場合、リクエストパケットを連続して転送する。このため、負荷分散装置P3とサーバP4との間で輻輳が発生し、リクエストパケットが消失することがあった。

【0098】

一方、本発明の第二実施形態における負荷分散装置3aは、複数のリクエストパケットで構成されるHTTPリクエストをサーバ4に転送する場合、先に転送されたリクエストパケットに対するAckを含むパケットを受信した後、後続のリクエストパケットを転送する。このため、負荷分散装置3aとサーバ4との間において、輻輳の発生が防止される。従って、輻輳発生によるリクエストパケットの消失が防止される。

【0099】

＜変形例＞

パケット送信許可部20は、送信を許可するリクエストパケットの数を指数関数的に増加させても良い。即ち、パケット送信許可部20は、送信を許可するリクエストパケットの数を、スロースタートメカニズムに従って決定しても良い。

【0100】

〔第三実施形態〕

＜システム構成＞

第三実施形態における負荷分散装置3bを用いた負荷分散システム1bについて説明する。負荷分散システム1bのシステム構成は、基本的に、第一実施形態における負荷分散装置3を用いた負荷分散システム1と同じである。即ち、負荷分散システム1bは、複数のクライアント2、負荷分散装置3b、複数のサーバ4、及び複数のクライアント2と負荷分散装置3bとを通信可能に接続するネットワーク5とを用いて構成される。クライアント2とサーバ4とは、上記した第一実施形態の欄ですでに説明したため、以下では負荷分散装置3bの構成について説明する。

【0101】

＜＜負荷分散装置＞＞

図17は、本発明による第三実施形態の負荷分散装置3bのブロック図である。

図17を用いて、負荷分散装置3bの構成について説明する。ただし、負荷分散装置3bについて、負荷分散装置3と異なる構成についてのみ説明する。負荷分散装置3bは、ペ

イロード有無確認部21をさらに備える。

【0102】

ペイロード有無確認部21は、CPUやRAM等、あるいは専用バケット処理プロセッサ等を用いて構成される。ペイロード有無確認部21は、バケットバッファ9に記憶されているバケットがペイロードを含むか否かを確認する。このバケットにペイロードが含まれていない場合、ペイロード有無確認部21は、このバケットを消去するように、バケットバッファ9に指示を出す。一方、このバケットにペイロードが含まれている場合、ペイロード有無確認部21は、このバケットをクライアント2へ転送するようにバケットバッファ9に指示を出す。

【0103】

〈動作例〉

図18は、第三実施形態における負荷分散装置3bの動作例のうち、負荷分散装置3と異なる部分のみを示した図である。図18を用いて、負荷分散装置3bの動作例について説明する。ただし、負荷分散装置3の動作と異なる動作についてのみ説明する。

【0104】

サーバ4から新たなバケットが受信されると、Ack番号抽出部18は、バケットバッファ8において、このバケットのAck番号よりもシーケンス番号及びデータ長の和が小さな値となるバケットの有無を確認する(S18)。このようなバケットがバケットバッファ8にバッファリングされている場合(S18-Yes)、バケットバッファ8は、そのバケットを消去する(S19)。

【0105】

次に、ペイロード有無確認部21は、このバケットがペイロードを含むか否かを判断する(S26)。このバケットがペイロードを含まない場合(S26-No)、ペイロード有無確認部21は、バケットバッファ9に、このバケットをクライアント2に転送せずに消去するように指示を出す。一方、このバケットがペイロードを含む場合(S26-Yes)、ペイロード有無確認部21は、このバケットをクライアント2へ転送するようにバケットバッファ9に指示を出す。

【0106】

〈動作シーケンス〉

図19は、第三実施形態における負荷分散装置3bを用いた負荷分散システム1bの動作シーケンスを示す図である。図19を用いて、負荷分散システム1bの動作シーケンスについて説明する。ただし、負荷分散システム1の動作シーケンスと異なる動作シーケンスについてのみ説明する。

【0107】

負荷分散装置3bは、サーバ4に対し、リクエストバケットa、bを転送する。サーバ4は、リクエストバケットaに対するAckを含むバケットや、リクエストバケットbに対するAckを含むバケット(この場合はレスポンスバケット)を、負荷分散装置3bに対し送信する。負荷分散装置3bは、リクエストバケットaに対するAckを含むバケットを受信した場合、このAckバケットをクライアント2に転送せずに消去する。即ち、負荷分散装置3bは、HTTPレスポンスを含まないバケットを受信した場合、このバケットをクライアント2に転送せずに消去する。一方、負荷分散装置3bは、レスポンスバケットを受信した場合、このレスポンスバケットをクライアント2に転送する。

【0108】

〈作用〉

図20は、従来技術における負荷分散装置P3と本発明の第三実施形態における負荷分散装置3bとの動作の差異を示す図である。図20を用いて、本発明の第三実施形態における負荷分散装置3bの作用について説明する。従来技術における負荷分散装置P3は、サーバP4から受信する全てのバケットをクライアントP2へ転送する。よって、従来技術における負荷分散装置P3は、HTTPレスポンスを含まないバケット(この場合、Ackのためのバケット)、即ちクライアントP2に転送される必要のないバケットをクライ

10

20

30

40

50

アントP2に転送する。このため、クライアントP2と負荷分散装置P3とを接続するネットワークの帯域が不必要に消費されていた。

【0109】

一方、本発明の第三実施形態における負荷分散装置3bは、HTTPレスポンスを含まないパケット、即ちペイロードを含まないパケットを、クライアント2に転送せずに消去する。このため、ネットワーク5において、帯域が無駄に消費されない。

【0110】

〈変形例〉

ペイロード有無確認部21は、本発明の第二実施形態における負荷分散装置3aに適用されても良い。本発明の第二実施形態における負荷分散装置3aでは、リクエストパケットaの転送とリクエストパケットbの転送との間に、必ずサーバ4からAckを含むパケット（ペイロードを含まないパケット）を受信する。このとき、ペイロード有無確認部21が適用されることにより、このパケットをクライアント2に送信することがなくなる。このため、ネットワーク5において、帯域を節約することが可能となる。

【0111】

〔その他〕

本発明は、以下のように特定することができる。

（付記1） クライアントから受信したデータの内容に応じて、クライアントへ当該データの確認応答を送信するか否かを判断する判断手段と、
前記判断手段の判断結果に応じて、クライアントへ確認応答を送信する確認応答手段と、
前記クライアントから受信したデータを複数のサーバのいずれかへ転送する転送手段と、
前記サーバへ転送されたデータのうち、このデータについてクライアントへ確認応答が送信されたデータのみを、前記サーバへの再送に備えて記憶する再送データ記憶手段と、
を備える負荷分散装置。

（付記2） 前記判断手段は、クライアントから受信したデータがデリミタを含まない場合にクライアントへ確認応答を送信すると判断し、前記データがデリミタを含む場合にクライアントへ確認応答を送信しないと判断する付記1に記載の負荷分散装置。

（付記3） 連続する複数のデータをサーバへ転送する場合に、最後にサーバに転送されたデータに対する確認応答をサーバから受信したときに、次のデータをサーバへ転送することを許可する確認応答判断手段をさらに備える付記1又は2に記載の負荷分散装置。

（付記4） サーバから受信されたデータを、クライアントへ転送すべきか否かを判断する転送判断手段をさらに備える付記1～3のいずれかに記載の負荷分散装置。

（付記5） 前記転送判断手段は、前記サーバから受信されたデータが、確認応答のためのデータのみからなる場合に転送すべきでないと判断する付記4に記載の負荷分散装置。

（付記6） 前記判断手段は、前記クライアントから受信したデータが一つのパケットからなるHTTPリクエストである場合に確認応答を送信すると判断する付記1に記載の負荷分散装置。

（付記7） 前記判断手段は、前記クライアントから受信したデータが複数のパケットからなるHTTPリクエストである場合に、このHTTPリクエストを構成する最後のパケットを受信した際に確認応答を送信すると判断する付記1又は2に記載の負荷分散装置。

【0112】

〔発明の効果〕

本発明によれば、クライアントからサーバへ転送するデータを、早期に送信バッファから消去することにより、送信バッファの使用時間を減少させることが可能となる。

〔図面の簡単な説明〕

〔図1〕 従来技術における再送管理省略の動作シーケンスを示す図である。

〔図2〕 スロースタートメカニズムを示す図である。

〔図3〕 従来技術の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。

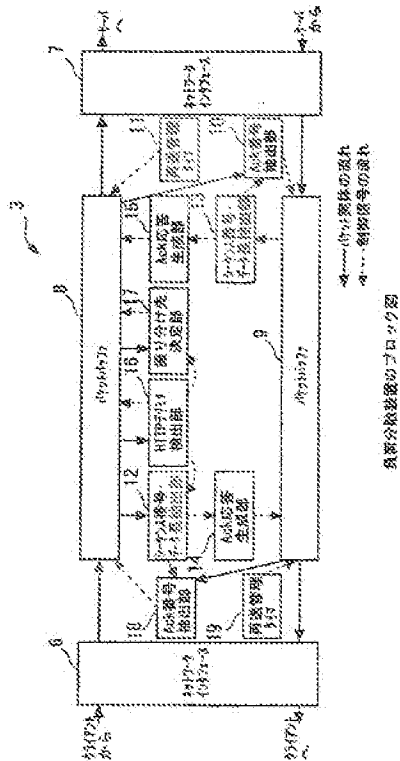
〔図4〕 負荷分散システムの概略を示す図である。

- 【図5】 第一実施形態の負荷分散装置のブロック図である。
- 【図6】 第一実施形態の負荷分散装置の動作例を示すフローチャートである。
- 【図7】 第一実施形態の負荷分散装置の動作例を示すフローチャートである。
- 【図8】 第一実施形態の負荷分散装置の動作例を示すフローチャートである。
- 【図9】 第一実施形態の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。
- 【図10】 第一実施形態の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。
- 【図11】 第一実施形態の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。 10
- 【図12】 負荷分散装置の記憶領域が解放されるタイミングを示す図である。
- 【図13】 第二実施形態の負荷分散装置のブロック図である。
- 【図14】 第二実施形態の負荷分散装置の動作例を示すフローチャートである。
- 【図15】 第二実施形態の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。
- 【図16】 従来技術における負荷分散装置と本発明の第二実施形態における負荷分散装置との動作の差異を示す図である。
- 【図17】 第三実施形態の負荷分散装置のブロック図である。
- 【図18】 第三実施形態の負荷分散装置の動作例を示すフローチャートである。
- 【図19】 第三実施形態の負荷分散装置を用いた負荷分散システムの動作シーケンスを示す図である。 20
- 【図20】 従来技術における負荷分散装置と本発明の第三実施形態における負荷分散装置との動作の差異を示す図である。

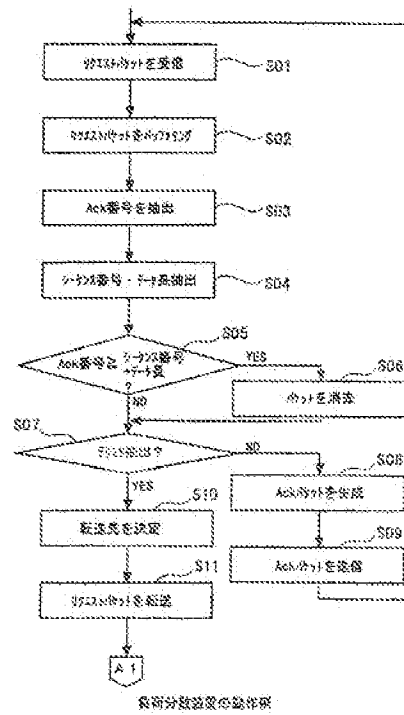
【符号の説明】

- | | |
|-------------|-----------------|
| 1 | 負荷分散システム |
| 2 | クライアント |
| 3, 3 a, 3 b | 負荷分散装置 |
| 4 | サーバ |
| 5 | ネットワーク |
| 6, 7 | ネットワークインタフェース |
| 8, 9 | パケットバッファ |
| 10, 18 | A c k番号抽出部 |
| 11, 19 | 再送管理タイマ |
| 12, 13 | シーケンス番号・データ長抽出部 |
| 14, 15 | A c k応答生成部 |
| 16 | H T T Pデリミタ検出部 |
| 17 | 振り分け先決定部 |
| 20 | パケット送信許可部 |
| 21 | ペイロード有無確認部 |

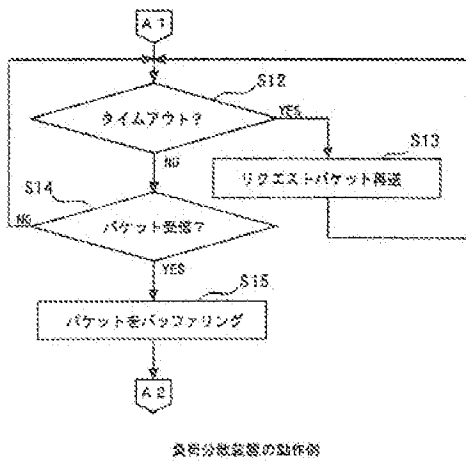
【図5】



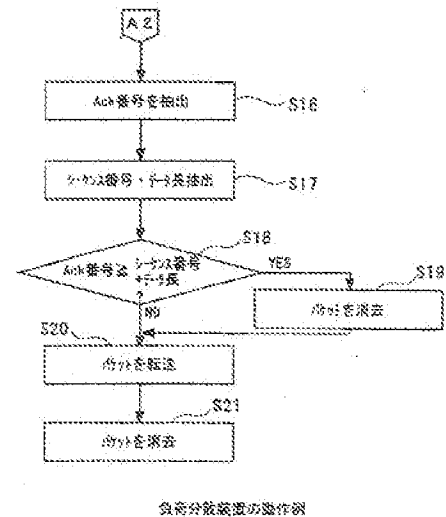
【図6】



【図7】

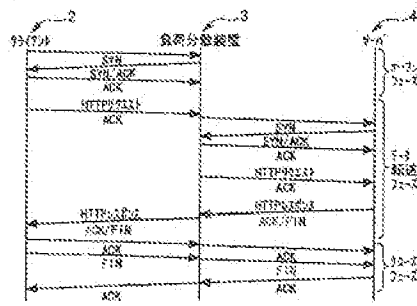


【図8】

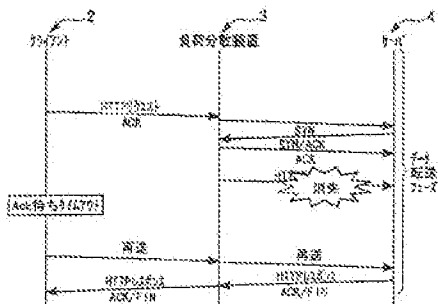


【図 9】

(a) パケットが消失しない場合

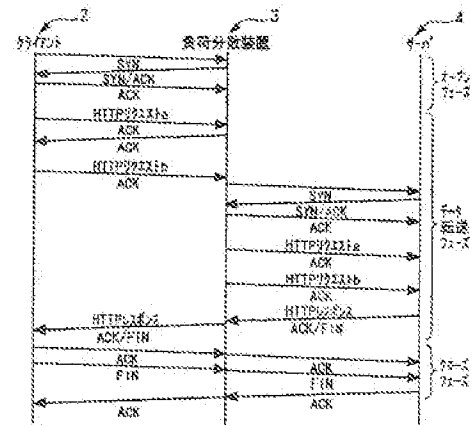


(b) パケットが消失する場合



負荷分散システムの動作シーケンス

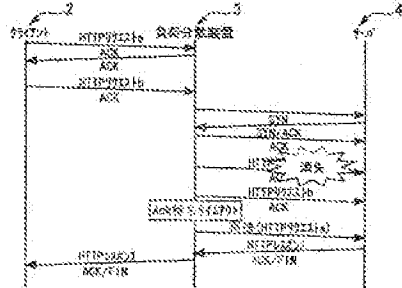
【図 10】



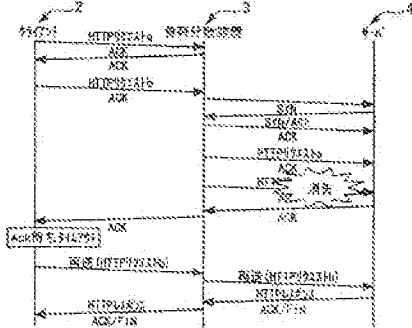
負荷分散システムの動作シーケンス

【図 11】

(a) リクエストパケットが消失する場合



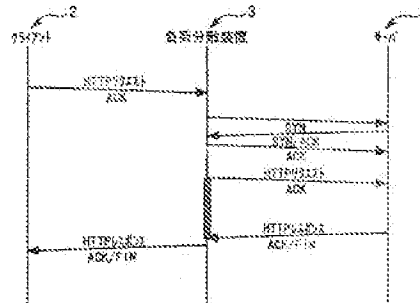
(b) リクエストパケットが消失する場合



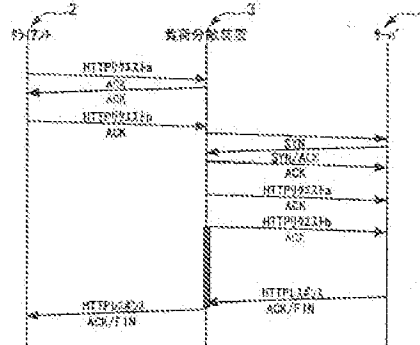
負荷分散システムの動作シーケンス

【図 12】

(a) HTTPリクエストが一つのパケットで構成されている場合

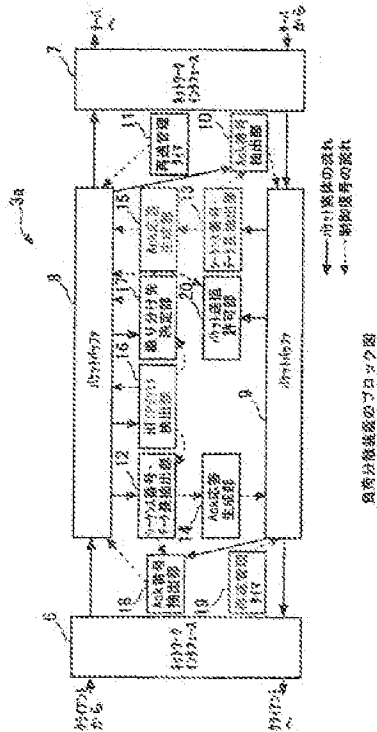


(b) HTTPリクエストが複数パケットで構成されている場合

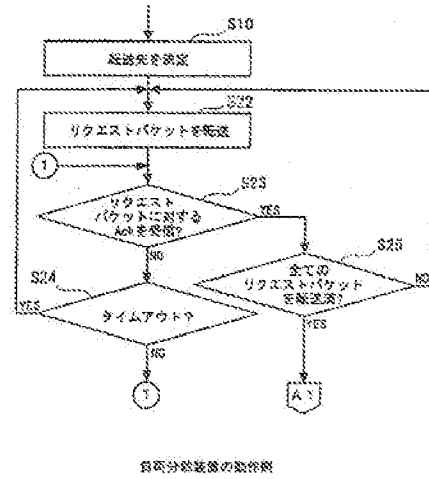


記憶領域が開放されるタイミング

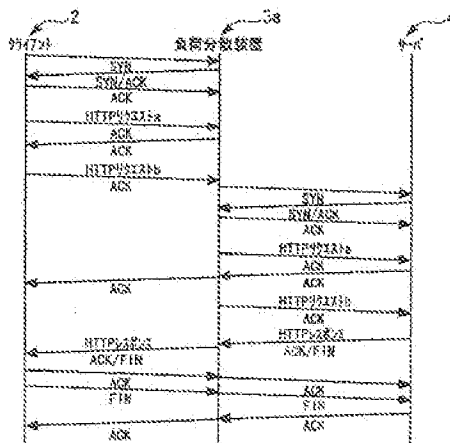
【図13】



【図14】

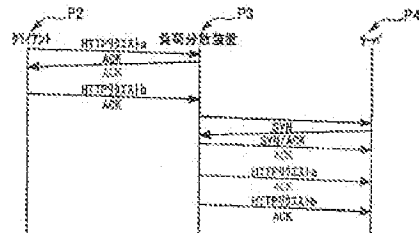


【図15】

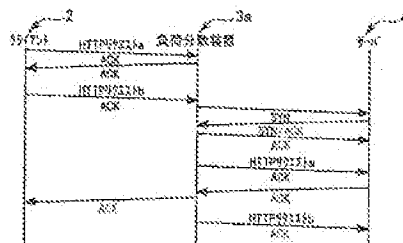


【図16】

(a) 従来技術の場合



(b) 本発明の第二実施形態の場合



【手続補正書】

【提出日】平成15年8月27日(2003.8.27)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0111

【補正方法】変更

【補正の内容】

【0111】

〔その他〕

本発明は、以下のように特定することができる。

(付記1) クライアントから受信したデータの内容に応じて、クライアントへ当該データの確認応答を送信するか否かを判断する判断手段と、
前記判断手段の判断結果に応じて、クライアントへ確認応答を送信する確認応答手段と、
前記クライアントから受信したデータを複数のサーバのいずれかへ転送する転送手段と、
前記サーバへ転送されたデータのうち、このデータについてクライアントへ確認応答が送信されたデータのみを、前記サーバへの再送に備えて記憶する再送データ記憶手段と、
を備える負荷分散装置。

(付記2) 前記判断手段は、クライアントから受信したデータがデリミタを含まない場合にクライアントへ確認応答を送信すると判断し、前記データがデリミタを含む場合にクライアントへ確認応答を送信しないと判断する付記1に記載の負荷分散装置。

(付記3) 連続する複数のデータをサーバへ転送する場合に、最後にサーバに転送されたデータに対する確認応答をサーバから受信したときに、次のデータをサーバへ転送することを許可する確認応答判断手段をさらに備える付記1又は2に記載の負荷分散装置。

(付記4) サーバから受信されたデータを、クライアントへ転送すべきか否かを判断する転送判断手段をさらに備える付記1～3のいずれかに記載の負荷分散装置。

(付記5) 前記転送判断手段は、前記サーバから受信されたデータが、確認応答のためのデータのみからなる場合に転送すべきでないと判断する付記4に記載の負荷分散装置。

(付記6) 前記判断手段は、前記クライアントから受信したデータが一つのケットからなるHTTPリクエストである場合に確認応答を送信しないと判断する付記1に記載の負荷分散装置。

(付記7) 前記判断手段は、前記クライアントから受信したデータが複数のケットからなるHTTPリクエストである場合に、このHTTPリクエストを構成する最後のケットを受信した際に確認応答を送信しないと判断する付記1又は2に記載の負荷分散装置。

フロントページの続き

Fターム(参考) 5K014 AA01 CA02 FA03

5K030 HA08 KA03 LA02 LE03 MB11 MB13

